

# Positive volatility simulation in the Heston model

Simon J.A. Malham · Anke Wiese

*Dedicated to Mrs Marlies Wiese*

Received: 22nd June 2010

**Abstract** In the Heston stochastic volatility model, the transition probability of the variance process can be represented by a non-central chi-square density. We focus on the case when the number of degrees of freedom is small and the zero boundary is attracting and attainable, typical in foreign exchange markets. We prove a new representation for this density based on sums of powers of generalized Gaussian random variables. Further we prove Marsaglia's polar method extends to this distribution, providing an exact method for generalized Gaussian sampling. The advantages are that for the mean-reverting square-root process in the Heston model and Cox-Ingersoll-Ross model, we can generate samples from the true transition density simply, efficiently and robustly.

**Keywords** stochastic volatility · positivity preservation · generalized Gaussian · generalized Marsaglia method

**Mathematics Subject Classification (2000)** 60H10 · 60H35 · 93E20

## 1 Introduction

A popular choice to model the evolution of an asset price and its stochastic volatility is the correlated Heston model [23]:

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{V_t} S_t (\rho dW_t^1 + \sqrt{1 - \rho^2} dW_t^2), \\ dV_t &= \kappa(\theta - V_t) dt + \varepsilon \sqrt{V_t} dW_t^1. \end{aligned}$$

Here  $(W^1, W^2)$  is a standard two-dimensional Wiener process, the parameters  $\mu$ ,  $\kappa$ ,  $\theta$  and  $\varepsilon$  are positive constants, and  $\rho \in (-1, 1)$ . The process  $S$  represents the price

---

Simon J.A. Malham · Anke Wiese  
 Maxwell Institute for Mathematical Sciences  
 and School of Mathematical and Computer Sciences  
 Heriot-Watt University, Edinburgh EH14 4AS, UK  
 Tel.: +44-131-4513200  
 Fax: +44-131-4513249  
 E-mail: S.J.Malham@ma.hw.ac.uk  
 E-mail: A.Wiese@ma.hw.ac.uk

process of an underlying asset, for example a stock index or an exchange rate, and  $V$  is the variance process of the log returns of  $S$ , modelled as a mean-reverting square-root process, the well-known Cox–Ingersoll–Ross process.

The transition probability density of the variance process  $V$  is known explicitly, it can be represented by a non-central chi-square density. Depending on the number of degrees of freedom  $\nu := 4\kappa\theta/\varepsilon^2$ , there are fundamental differences in the behaviour of the variance process. If  $\nu$  is larger or equal to 2, the zero boundary is unattainable; if it is smaller than 2, the zero boundary is attracting and attainable. At the zero boundary though, the solution is immediately reflected into the positive domain.

A number of successful simulation schemes have been developed for the non-attainable zero boundary case. There are schemes based on implicit time-stepping integrators, see for example Alfonsi [3, 4] and Kahl and Schurz [33]. Other time discretization approaches involve splitting the drift and diffusion vector fields and evaluating their separate flows (sometimes exactly) before they are recomposed together, typically using the Strang ansatz. See for example Higham, Mao and Stuart [26], Ninomiya and Victoir [45] and Alfonsi [4]. However, the splitting methods and the implicit methods so far only apply in the non-attracting zero boundary case.

Other direct discretization approaches, that can be applied to the attainable and unattainable zero boundary case are based on forced Euler–Maruyama approximations and typically involve negativity truncations; some of these methods are positivity preserving. See for example Deelstra and Delbaen [14], Bossy and Diop [9] and also Berkaoui, Bossy and Diop [7], Lord, Koekkoek and Van Dijk [39], as well as Higham and Mao [25], among others. These methods all converge to the exact solution, but their rate of strong convergence is difficult to establish. However, this said, the leading method in this class—the full truncation method of Lord, Koekkoek and Van Dijk [39]—has in practice proved highly effective.

Exact simulation methods typically sample from the known non-central chi-square distribution  $\chi_\nu^2(\lambda)$  for the transition probability of the variance process  $V$  (see Cox, Ingersoll and Ross [13] and Glassermann [19, Section 3.4]). Broadie and Kaya [10] proposed sampling from  $\chi_\nu^2(\lambda)$  as follows. When  $\nu > 1$ ,  $\chi_\nu^2(\lambda) = (N(0, \sqrt{\lambda}))^2 + \chi_{\nu-1}^2$ , so such a sample can be generated by a standard Normal sample and a central chi-square sample. When  $0 < \nu < 1$ , such a sample can be generated by sampling from a Poisson distribution with mean  $\lambda/2$ , and then sampling from a central  $\chi_{2N+\nu}^2$  distribution. To simulate the asset price Broadie and Kaya integrated the volatility process to obtain an expression for  $\int \sqrt{V_\tau} dW_\tau$ . They substituted that expression into the stochastic differential equation for  $\ln S_t$ . The most difficult task left is then to simulate  $\int V_\tau d\tau$  on the global interval of integration conditioned on the endpoint values of  $V$ ; see Smith [53]. The Laplace transform of the transition density for this integral is known from results in Pitman and Yor [47]. Broadie and Kaya used Fourier inversion techniques to sample from this transition density. Glasserman and Kim [20] on the other hand, showed that linear combinations of series of particular gamma random variables exactly sample this density. They used truncations of those series to generate suitably accurate sample approximations. Anderson [5] suggested two approximations that make simulation of the Heston model very efficient. The first was, after discretizing the time interval of integration for the price process, to approximate  $\int V_\tau d\tau$  on the integration subinterval by a simple quadrature. This would thus require non-central  $\chi_\nu^2(\lambda)$  samples for the volatility at each timestep. Hence the second, was to approximate and thus efficiently sample, the  $\chi_\nu^2(\lambda)$  distribution—in two different ways depending on the size of  $\lambda$ . Haastrecht

and Pelsser [22] have recently introduced a rival  $\chi_\nu^2(\lambda)$  sampling method to Andersen's. Moro and Schurz [44] have also successfully combined exponential splitting with exact simulation. Dyrting [15] outlines and compares several different series and asymptotic approximations for non-central chi-square distribution.

There are also numerous approximation methods based on the corresponding Fokker–Planck partial differential equation. These can take the form of Fourier transform methods—see Carr and Madan [11], Kahl and Jäckel [32] or Fang and Oosterlee [16, 17] for example—or some involve direct discretization of the Fokker–Planck equation.

We focus on the challenge of the attainable zero boundary case and in particular on the case when  $\nu \ll 1$ , typical of FX markets and long-dated interest rate markets as remarked in Andersen [5]. The method we propose follows the lead of Andersen [5]. It is based on efficiently simulating the known non-central  $\chi_\nu^2(\lambda)$  transition density for the volatility required for each timestep of Andersen's integration method for the Heston model. More precisely, we suggest an *exact simulation* method for the non-central  $\chi_\nu^2(\lambda)$  density as follows. A non-central  $\chi_\nu^2(\lambda)$  random variable can be generated from a central  $\chi_{2N+\nu}^2$  random variable with  $N$  chosen from a Poisson distribution with mean  $\lambda/2$ . Further, a  $\chi_{2N+\nu}^2$  random variable can be generated from the sum of squares of  $2N$  independent standard Normal random variables and an independent central  $\chi_\nu^2$  random variable. So the question we now face is how can we efficiently simulate a central  $\chi_\nu^2$  random variable, especially for  $\nu < 1$ ? Suppose that  $\nu$  is rational and expressed in the form  $\nu = p/q$  with  $p$  and  $q$  natural numbers. We show that a central  $\chi_\nu^2$  random variable can be generated from the sum of the  $2q$ th power of  $p$  independent random variables chosen from a generalized Gaussian distribution  $N(0, 1, 2q)$ , where a  $N(0, 1, q)$  distribution has density

$$f_{N(0,1,q)}(x) := \frac{q}{2^{1+1/q}\Gamma(1/q)} \cdot \exp\left(-\frac{1}{2}|x|^q\right),$$

where  $x \in \mathbb{R}$  and  $\Gamma(\cdot)$  is the standard gamma function. How can we sample from a  $N(0, 1, 2q)$  distribution? Our answer lies in generalizing Marsaglia's polar method for pairs of independent standard Normal random variables. Indeed we generate  $2q$  uniform random variables  $U = (U_1, \dots, U_{2q})$  over  $[-1, 1]$ , and condition on their  $2q$ th norm  $\|U\|_{2q}$ , being less than unity. Then we prove that the  $2q$  random variables  $U \cdot (-2 \log \|U\|_{2q}^{2q})^{1/2q} / \|U\|_{2q}$  are independent  $N(0, 1, 2q)$  random variables. We provide a thorough comparison, of our generalized Marsaglia polar method for sampling from the central  $\chi_\nu^2$  distribution, to the acceptance-rejection method of Ahrens and Dieter (see Ahrens and Dieter [2] and Glasserman [19]).

The Cox–Ingersoll–Ross process, which has a non-central chi-squared transition probability, can thus be exactly simulated by the approach just described, which we will call the Marsaglia generalized Gaussian method (MAGG). The *advantages* of this approach are that for the mean-reverting variance process in the Heston model, we can generate high quality samples simply and robustly. The *disadvantage* of MAGG when simulating the Cox–Ingersoll–Ross process is that the degrees of freedom  $\nu$  must be rational (however, in practice, this is typically fulfilled). We demonstrate our method in the computation of option prices for parameter cases that are considered in Andersen [5] and Glasserman and Kim [20] and described there as challenging and practically relevant. To summarize, we:

- Prove that a central chi-squared random variable with less than one degree of freedom, can be written as a sum of powers of generalized Gaussian random variables;

- Prove a new method—the generalized Marsaglia polar method—for generating generalized Gaussian samples;
- Establish a new simple, exact, unbiased and efficient method for simulating the Cox–Ingersoll–Ross process, for an attracting and attainable zero boundary, and thus establish a new simple method for simulating the Heston model.

Our paper is organised as follows. In Section 2 we present our new method for sampling from the non-central chi-squared distribution based on sampling from the generalized Gaussian distribution. We include a thorough numerical comparison with the acceptance-rejection method of Ahrens and Dieter [2]. We apply our MAGG method to the Heston model in Section 3. We compare its accuracy and efficiency to the method of Andersen [5]. Finally in Section 4 we present some concluding remarks.

## 2 Non-central chi-square sampling

We present our new theoretical probabilistic results. We begin by introducing the generalized Gaussian distribution and show that central  $\chi_\nu^2$  random variables can be represented by sums of powers of generalized Gaussian random variables. To generate central  $\chi_\nu^2$  samples in this way, we require an efficient method for generating generalized Gaussian samples. We provide such a method in the form of a generalization of Marsaglia’s polar method for standard Normal random variables. Then, to put our approach on a firm practical footing, we compare its efficiency to the most well-known exact  $\chi_\nu^2$  sampling method, the acceptance-rejection method of Ahrens and Dieter [2]. Rounding off this section, we return to our original goal, and explicitly state our general algorithm for non-central  $\chi_\nu^2(\lambda)$ . This will be the algorithm we use in our application to the Heston model in Section 3.

### 2.1 Generalized Gaussian distribution

We prove that random variables with a central  $\chi_\nu^2$  distribution, especially for  $\nu < 1$ , can be represented by random variables with a generalized Gaussian distribution.

**Definition 1 (Generalized Gaussian distribution)** A generalized  $N(0, 1, q)$  random variable, for  $q \geq 1$ , has density

$$f_{N(0,1,q)}(x) := \frac{q}{2^{1/q+1}\Gamma(1/q)} \cdot \exp(-\tfrac{1}{2}|x|^q),$$

where  $x \in \mathbb{R}$  and  $\Gamma(\cdot)$  is the standard gamma function.

See Gupta and Song [21], Song and Gupta [54], Sinz, Gerwinn and Bethge [52] and Sinz and Bethge [51] for more details on this distribution and its properties.

**Theorem 1 (Central chi-square from generalized Gaussians)** Suppose  $X_i \sim N(0, 1, 2q)$  are independent identically distributed random variables for  $i = 1, \dots, p$ , where  $q \geq 1$  and  $p \in \mathbb{N}$ . Then we have

$$\sum_{i=1}^p |X_i|^{2q} \sim \chi_{p/q}^2.$$

*Proof* If  $X \sim N(0, 1, 2q)$ , then we see that

$$\begin{aligned} \mathbb{P}(|X|^{2q} < x) &= \frac{2q}{2^{1/2q} \Gamma(1/2q)} \int_0^{|x|^{1/2q}} \exp\left(-\frac{1}{2}|\tau|^{2q}\right) d\tau \\ &= \frac{1}{2^{1/2q} \Gamma(1/2q)} \int_0^x \xi^{1/2q-1} \exp\left(-\frac{1}{2}\xi\right) d\xi \end{aligned}$$

where  $\xi = |\tau|^{2q}$ . Hence we deduce that  $|X|^{2q} \sim \chi_{1/q}^2$ . Now using that the sum of  $p$  independent identically distributed  $\chi_{1/q}^2$  random variables have a  $\chi_{p/q}^2$  distribution establishes the result.  $\square$

## 2.2 Generalized Marsaglia polar method

If we intend to use  $N(0, 1, 2q)$  samples to generate  $\chi_{p/q}^2$  samples, we need an accurate and efficient method for sampling from a generalized Gaussian distribution. To this end we generalize Marsaglia's polar method for pairs of independent standard Normal random variables (see Marsaglia [42]).

**Theorem 2 (Generalized Marsaglia polar method)** *Suppose for some  $q \in \mathbb{N}$  that  $U_1, \dots, U_q$  are independent identically distributed uniform random variables over  $[-1, 1]$ . Condition this sample set to satisfy the requirement  $\|U\|_q < 1$ , where  $\|U\|_q$  is the  $q$ -norm of  $U = (U_1, \dots, U_q)$ . Then the  $q$  random variables generated by  $U \cdot (-2 \log \|U\|_q^q)^{1/q} / \|U\|_q$  are independent  $N(0, 1, q)$  distributed random variables.*

*Proof* Suppose for some  $q \in \mathbb{N}$  that  $U = (U_1, \dots, U_q)$  are independent identically distributed uniform random variables over  $[-1, 1]$ , conditioned on the requirement that  $\|U\|_q < 1$ . Then the scalar variable

$$Z := (-2 \log \|U\|_q^q)^{1/q} > 0$$

is well defined. Let  $f$  denote the probability density function of  $U$  given  $\|U\|_q < 1$ ; it is defined on the interior of the  $q$ -sphere,  $\mathbb{S}_q(1)$ , whose bounding surface is  $\|U\|_q = 1$ . We define a new set of  $q$  random variables  $W = (W_1, \dots, W_q)$  by the map  $G: \mathbb{S}_q(1) \rightarrow \mathbb{R}^p$  where  $G: U \mapsto W$  is given by

$$G \circ U = \frac{Z}{\|U\|_q} \cdot U.$$

Note that the inverse map  $G^{-1}: \mathbb{R}^p \rightarrow \mathbb{S}_q(1)$  is well defined and given by

$$G^{-1} \circ W = \frac{\exp(-Z^q/2q)}{Z} \cdot W,$$

where we note that in fact  $Z = \|W\|_q$  which comes from taking the  $q$ -norm on each side of the relation  $W = G(U)$ .

We wish to determine the probability density function of  $W$ . Note that if  $\Omega \subset \mathbb{R}^q$ ,

$$\begin{aligned} \mathbb{P}(W \in \Omega) &= \mathbb{P}(U \in G^{-1}(\Omega)) \\ &= \int_{G^{-1}(\Omega)} f \circ u \, du \\ &= \int_{\Omega} (f \circ G^{-1} \circ w) \cdot |\det(DG^{-1} \circ w)| \, dw, \end{aligned}$$

where for  $w = (w_1, \dots, w_p) \in \Omega$ , the quantity  $DG^{-1} \circ w$  denotes the Jacobian transformation matrix of  $G^{-1}$ . Hence the probability density function of  $W$  is given by

$$(f \circ G^{-1} \circ w) \cdot |\det(DG^{-1} \circ w)|.$$

The Jacobian matrix and its determinant are established by direct computation. For each  $i, k = 1, \dots, q$  we see that if we define  $g(z) := -(1/2 + 1/z^q)$ , then

$$\frac{\partial G_k^{-1}}{\partial w_i} = \frac{\exp(-z^q/2q)}{z} \cdot \left( \delta_{ik} + g(z) \cdot (\operatorname{sgn}(w_i) \cdot |w_i|^{q-1}) \cdot w_k \right),$$

where  $\delta_{ik}$  is the Kronecker delta function. If we set

$$v = (\operatorname{sgn}(w_1) \cdot |w_1|^{q-1}, \dots, \operatorname{sgn}(w_q) \cdot |w_q|^{q-1})^T$$

then we see that our last expression generates the following relation for the Jacobian matrix (here  $I_q$  denotes the  $q \times q$  identity matrix):

$$\frac{z}{\exp(-z^q/2q)} \cdot (DG^{-1} \circ w) = I_q + g(z) \cdot v w^T.$$

From the determinant rule for rank-one updates—see Meyer [43, p. 475]—we see that the determinant of the Jacobian matrix is given by

$$\begin{aligned} \det(DG^{-1} \circ w) &= \frac{\exp(-z^q/2)}{z^q} \cdot (1 + g(z) w^T v) \\ &= \frac{\exp(-z^q/2)}{z^q} \cdot (1 + g(z) z^q) \\ &= -\frac{1}{2} \exp(-z^q/2), \end{aligned}$$

where we used the definition for  $g(z)$  in the last step. Noting that  $\operatorname{vol}(\mathbb{S}_q(1)) = 2^q \cdot (\Gamma(1/q))^q / q^q$  we have

$$(f \circ G^{-1} \circ w) \cdot |\det(DG^{-1} \circ w)| = \frac{q^q}{2^{q+1} (\Gamma(1/q))^q} \cdot \exp(-z^q/2).$$

This is the joint probability density function for  $q$  independent identically distributed  $q$ -generalized Gaussian random variables, establishing the required result.  $\square$

### 2.3 Comparison with acceptance-rejection method

We compare our approach for generating central  $\chi_\nu^2$  samples to the acceptance-rejection algorithm of Ahrens and Dieter [2] for the case  $\nu < 2$ . In particular we use the form of this acceptance-rejection algorithm outlined in Glasserman [19, pp. 126–7]. The acceptance-rejection algorithm is based on a mixture of the prior densities  $(\nu/2) x^{\nu/2-1}$  on  $[0, 1]$  and  $\exp(1-x)$  on  $(1, \infty)$ , with weights  $e/(e + \nu/2)$  and  $(\nu/2)/(e + \nu/2)$ , respectively; here  $e = \exp(1)$ . This method generates one  $\chi_\nu^2$  random variable with probability of acceptance

$$P_{AD} := \frac{(\nu/2) \Gamma(\nu/2) e}{\nu/2 + e}.$$

In this method, the number of degrees of freedom  $\nu$  can be any real number.

For our generalized Marsaglia polar method, we restrict ourselves to case when the number of degrees of freedom is rational, i.e.  $\nu = p/q$  with  $p, q \in \mathbb{N}$ . The algorithm for generating central  $\chi_\nu^2$  samples is as follows.

**Algorithm 1 (Central chi-square samples)** To produce an exact  $\chi_{p/q}^2$  sample:

1. Generate  $2q$  independent uniform random variables over  $[-1, 1]$ :  $U = (U_1, \dots, U_{2q})$ .
2. If  $\|U\|_{2q} < 1$  continue, otherwise repeat Step 1.
3. Compute  $Z = U \cdot (-2 \log \|U\|_{2q}^{2q})^{1/2q} / \|U\|_{2q}$ . This gives  $2q$  independent  $N(0, 1, 2q)$  distributed random variables  $Z = (Z_1, \dots, Z_{2q})$ .
4. Compute  $Z_1^{2q} + \dots + Z_p^{2q}$ .

In the second step, the probability of accepting  $U_1, \dots, U_{2q}$  is given by the ratio of the volumes of  $\mathbb{S}_{2q}(1)$  and  $[-1, 1]^{2q}$ :

$$P_{\text{Mar}} := \left( \frac{\Gamma(1/2q)}{2q} \right)^{2q}.$$

Note for  $q = 1$ , the probability of acceptance is 0.7854. Further as  $q \rightarrow \infty$  we have  $P_{\text{Mar}} \rightarrow \exp(-\gamma) \approx 0.5615$ . Here  $\gamma$  is the Euler–Mascheroni constant and we have used that  $\Gamma(z) \sim 1/z - \gamma$  as  $z \rightarrow 0^+$ .

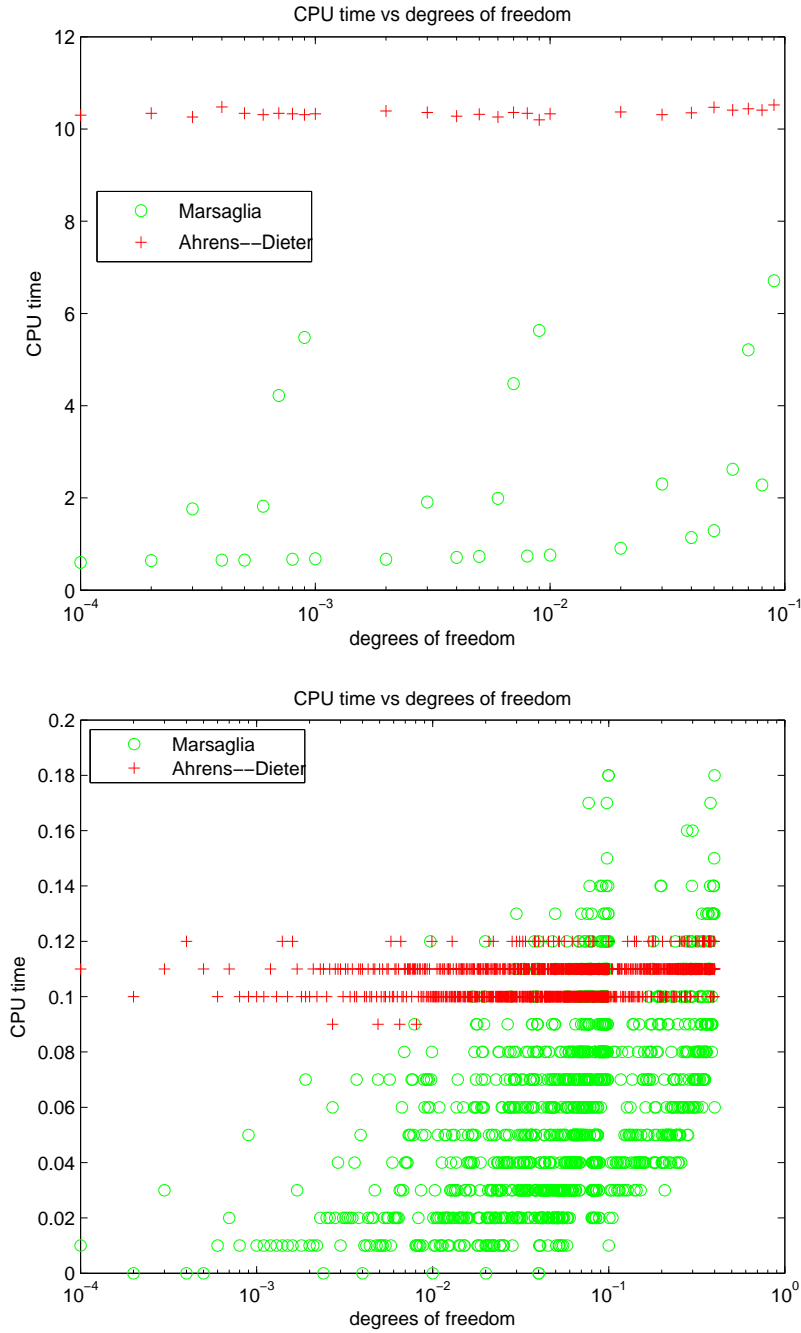
In practice we will need to generate a large number of samples. For the generalized Marsaglia polar method, in each accepted attempt, we generate  $2q$  generalized Gaussian random variables. Of these,  $p$  random variables are used to generate a  $\chi_{p/q}^2$  random variable. The number of attempts until the first success has a geometric distribution with mean  $1/P_{\text{Mar}}$ . Hence the expected number of steps to generate  $2q/p$  independent  $\chi_{p/q}^2$  random variables is thus  $1/P_{\text{Mar}}$ . For the acceptance-rejection method, the expected number of attempts to generate  $2q/p$  independent  $\chi_{p/q}^2$  distributed random variables is  $(2q/p) \cdot (1/P_{\text{AD}})$ .

The first natural question is whether the expected number of attempts for the generalized Marsaglia polar method is less than that for the acceptance-rejection method. In other words, to generate  $2q/p$  random variables, is  $1/P_{\text{Mar}} \leq (2q/p) \cdot (1/P_{\text{AD}})$ ? Or equivalently, when does  $p/q \leq 2P_{\text{Mar}}/P_{\text{AD}}$  hold? We examine the right-hand side more carefully; set  $z := 1/2q$ , so  $0 < z < 1/2$ . Then we have

$$\frac{P_{\text{Mar}}}{P_{\text{AD}}} = (z \Gamma(z))^z \cdot \frac{\nu/2 + e}{(\nu/2) \Gamma(\nu/2) e}.$$

Note  $z$  and  $\nu/2$  are independent. A lower bound for  $(z \Gamma(z))^z$  is  $\exp(-\gamma) \approx 0.5615$  for  $0 < z < 1/2$ , whilst a lower bound for  $(\nu/2 + e)/((\nu/2) \Gamma(\nu/2) e)$  is 1 for  $0 < \nu < 2$ . Hence  $2P_{\text{Mar}}/P_{\text{AD}} > 1$  and so for  $p/q < 1$ , the expected number of attempts for the generalized Marsaglia method is less than that for the acceptance-rejection method. We further note that the expected number of attempts for the generalized Marsaglia method to generate  $2q/p$  chi-square samples is bounded by its value for  $q = 1$  and the limit as  $q \rightarrow \infty$ , more precisely the expected number of attempts is  $1/P_{\text{Mar}} \in (1.2732, 1.7811)$ . In contrast in the acceptance-rejection method, the expected number of attempts to generate  $2q/p$  samples is  $(2q/p)(1/P_{\text{AD}})$ , which is unbounded.

The second natural question is how do the two algorithms perform in practice? The issue that immediately surfaces is that the generalized Marsaglia method is restricted to rational numbers. However, in practical applications this is not restrictive. In Figure 1 the upper panel shows the CPU time needed by both methods to generate  $10^6$  central  $\chi_\nu^2$  samples for the values  $\nu = n \cdot 10^{-m}$  where  $n = 1, \dots, 9$  and  $m = 2, 3, 4$ . We observe that the Ahrens and Dieter acceptance-rejection method roughly requires the same CPU time to generate central  $\chi_\nu^2$  samples for these values of  $\nu$ . It is slower



**Fig. 1** CPU time versus the number of degrees of freedom  $\nu$  for the generalized Marsaglia method and the Ahrens and Dieter acceptance-rejection method. The upper panel shows the CPU time needed to generate  $10^6$  samples for the values  $\nu = n \cdot 10^{-m}$  where  $n = 1, \dots, 9$  and  $m = 2, 3, 4$ . The lower panel shows the CPU time to need to generate  $10^4$  samples, simultaneously for two sets of  $\nu$  values given to three significant figures, namely  $\nu = (1 + m) \cdot 10^{-4}$  for  $m = 0, 1, \dots, 1000$  and  $\nu = 0.101 + m \cdot 10^{-3}$  for  $m = 0, 1, \dots, 299$ .



than for the generalized Marsaglia method. However the generalized Marsaglia method shows more variation in the CPU time required. In particular for example, for values of  $\nu$  equal to 3, 6, 7 and 9 times  $10^{-m}$  for all  $m = 2, 3, 4$ , it takes longer to generate central  $\chi_\nu^2$  samples than for the other  $\nu$  values. This is due to the fact that as rational numbers, with denominators as powers of 10, they do not simplify nicely to what might be considered the optimal format for sampling with this method, namely  $1/q$ . For values of  $\nu$  which cannot be reduced to this optimal format, we need to sum over a number of generalized Gaussian samples to produce a central  $\chi_\nu^2$  sample. However any decimal with a finite number of significant figures can be written as the sum of fractions of powers of 10. Further a central  $\chi_\nu^2$  random variable can be constructed by adding central  $\chi_{\nu_i}^2$  random variables for which  $\nu_1 + \dots + \nu_k = \nu$ . Suppose indeed, we generate a central  $\chi_\nu^2$  sample by adding  $\chi_{\nu_i}^2$  samples where the  $\nu_i$  are fractions of powers of 10 that generate each significant figure. From the upper panel in Figure 1, we observe that provided  $\nu$  does not have too many significant figures, then on average, the generalized Marsaglia method will be more efficient than the acceptance-rejection method. This is in fact confirmed in the lower panel in Figure 1. There we show how the CPU time varies with the number of degrees of freedom, when  $\nu$  is given to 3 significant figures. Even with the number of degrees of freedom given to 4 significant figures, we can see from the lower panel in Figure 1 that the generalized Marsaglia method will still be more efficient on average. Parameter values such as the number of degrees of freedom  $\nu$  are often determined by calibration. Since these are often quoted to only 2 or 3 significant figures, the generalized Marsaglia method would be the method of choice. We return to this issue in our conclusion section.

All simulations were run in Matlab, whose Profiler feature reveals that for the generalized Marsaglia method most CPU time is spent on computing  $\|U\|_{2q}$  in Step 2 and the sum in Step 4. In the acceptance-rejection method most time is spent on the decision processes required for choosing which of the mixture of prior densities to use.

## 2.4 Non-central chi-squared sampling

We return to *exact simulation* of the non-central  $\chi_\nu^2(\lambda)$  distribution. A  $\chi_\nu^2(\lambda)$  random variable can be generated as follows. Choose a random variable  $N$  from a Poisson distribution with mean  $\lambda/2$ . Then a sample generated from a central  $\chi_{2N+\nu}^2$  distribution is in fact a  $\chi_\nu^2(\lambda)$  sample. In other words we have

$$F_{\chi_\nu^2(\lambda)}(z) = \sum_{k=0}^{\infty} \mathbb{P}(N = k) \cdot F_{\chi_{2k+\nu}^2}(z).$$

See for example Johnson [29] or Brodie and Kaya [10]. Hence we are left with the problem of how to sample from a  $\chi_{2N+\nu}^2$  distribution. If we generate  $2N$  independent standard Normal random variables, say  $Y_1, \dots, Y_{2N}$ , and an independent  $\chi_\nu^2$  random variable, say  $Z$ , then  $Y_1^2 + \dots + Y_{2N}^2 + Z \sim \chi_{2N+\nu}^2$ . Putting all the components together we arrive at the following simple algorithm. We assume  $\nu = p/q$  with  $p$  and  $q$  natural numbers.

**Algorithm 2 (Non-central chi-square samples)** To produce an exact  $\chi_{p/q}^2(\lambda)$  sample:

1. Use Algorithm 1 to generate  $2q$  independent  $N(0, 1, 2q)$  distributed random variables  $Z = (Z_1, \dots, Z_{2q})$ .
2. Generate Poisson distributed random variable  $N$  with mean  $\lambda/2$ .
3. Generate  $2N$  standard Normal random variables, call them  $Y_1, \dots, Y_{2N}$ .
4. Compute  $Y_1^2 + \dots + Y_{2N}^2 + Z_1^{2q} + \dots + Z_p^{2q} \sim \chi_{p/q}^2(\lambda)$ .

Note that if  $p < 2q$  then we can use the remaining  $N(0, 1, 2q)$  random variables we generate in Step 3, the next time we need to generate a  $\chi_{p/q}^2(\lambda)$  sample. In practice we don't really need to consider the case  $p \geq 2q$ , but for the sake of completeness, we would simply generate  $p - 2q$  more  $N(0, 1, 2q)$  samples by repeating Steps 1–3.

*Remark 1 (Poisson sampling)* As Haastrecht and Pelsner [22] note, the mean  $\mu = \lambda/2$  of the Poisson distribution we wish to sample from is small. We can efficiently draw a sample from such a Poisson distribution by inverting its distribution function over a uniform random variable—see Knuth [36, p. 137] or Ahrens and Dieter [1]. The algorithm is as follows. Calculate  $\exp(-\mu)$ . Generate uniform random variables (over  $[0, 1]$ ) say  $U_1, U_2, \dots$  until  $U_1 \cdot U_2 \cdots U_m \leq \exp(-\mu)$ . Set  $N \leftarrow m - 1$ . On average this algorithm requires the generation of  $\mu + 1$  uniform variates.

### 3 Application: the Heston model

The *Heston model* (Heston [23]) is a two-factor model, in which the first component  $S$  describes the evolution of a financial variable such as a stock index or exchange rate, and the second component  $V$  describes the stochastic variance of its returns. The Heston model is given by

$$\begin{aligned} dS_t &= \mu S_t dt + \rho \sqrt{V_t} S_t dW_t^1 + \sqrt{1 - \rho^2} \sqrt{V_t} S_t dW_t^2, \\ dV_t &= \kappa(\theta - V_t) dt + \varepsilon \sqrt{V_t} dW_t^1, \end{aligned}$$

where  $W_t^1$  and  $W_t^2$  are independent scalar Wiener processes. The parameters  $\mu, \kappa, \theta$  and  $\varepsilon$  are all positive and  $\rho \in (-1, 1)$ . In the context of option pricing, a pricing measure must be specified. We assume here that the dynamics of  $S$  and  $V$  as specified above are given under the pricing measure. For a discussion and derivation of various equivalent martingale measures in the Heston model see for example Hobson [24]. By the Yamada condition this model has a unique strong solution. In particular, the volatility  $V$  is non-negative, and the stock price  $S$ , as a pure exponential process, is positive. Without loss of generality we suppose  $\mu = 0$ . We explain how we simulate the variance process  $V$  for the case of an attracting and attainable zero boundary first. Then we discuss how we approximately simulate the asset price process  $S$ . Finally we present some explicit simulation results.

#### 3.1 Variance process simulation

The variance process  $V_t$ , generated by the scalar stochastic differential equation above, is modelled as a mean-reverting process with mean  $\theta$ , rate of convergence  $\kappa$  and square root diffusion scaled by  $\varepsilon$ . It is known as the *Cox–Ingersoll–Ross process* (see Cox,

Ingersoll and Ross [13] who modelled the short rate of interest using this process). We define the *degrees of freedom* for this process to be

$$\nu := 4\kappa\theta/\varepsilon^2.$$

When  $\nu \in \mathbb{N}$  the process  $V_t$  can be reconstructed from the sum of squares of  $\nu$  Ornstein–Uhlenbeck processes; hence the label of degrees of freedom. When  $\nu < 2$  the zero boundary is attracting and attainable, while when  $\nu \geq 2$ , the zero boundary is non-attracting. These properties are immediate from the Feller boundary criteria, see Feller [18]. These are based on inverting the associated stationary elliptic Fokker–Planck operator, with boundary conditions, and can be found for example in Karlin and Taylor [35].

Here we focus on the challenge of  $\nu < 2$  and in particular cases when  $\nu \ll 1$  typical of FX markets (Andersen [5]). Importantly, though the zero boundary is attracting and attainable, it is strongly reflecting—if the process reaches zero it leaves it immediately and bounces back into the positive domain—see Revuz and Yor [48, p. 412]. We detailed in the introduction how this case is a major obstacle, particularly for direct discretization methods. A comprehensive account of direct discretization methods can be found in Lord, Koekkoek and Van Dijk [39], to where the reader is referred. Based on our experience, the *full truncation method* proposed by Lord, Koekkoek and Van Dijk has so far proven to be the most accurate and efficient of this class. This is also evidenced by Andersen [5] and Haastrecht and Pelsser [22] who complete thorough comparisons.

The method we propose follows the lead of Broadie and Kaya [10] and Andersen [5], and is based on simulating the known transition probability density for the Cox–Ingersoll–Ross process. We quote the following form for this transition density, that can be found in Cox, Ingersoll and Ross [13], from a proposition in Andersen [5].

**Proposition 1** *Let  $F_{\chi_\nu^2(\lambda)}(z)$  be the cumulative distribution function for the non-central chi-squared distribution with  $\nu$  degrees of freedom and non-centrality parameter  $\lambda$ :*

$$F_{\chi_\nu^2(\lambda)}(z) = \frac{\exp(-\lambda/2)}{2^{\nu/2}} \sum_{j=0}^{\infty} \frac{(\lambda/2)^j}{j! 2^j \Gamma(\nu/2 + j)} \int_0^z \xi^{\nu/2+j-1} \exp(-\xi/2) d\xi.$$

Set  $\nu := 4\kappa\theta/\varepsilon^2$  and define

$$\eta(h) := \frac{4\kappa \exp(-\kappa h)}{\varepsilon^2 (1 - \exp(-\kappa h))},$$

where  $h = t_{n+1} - t_n$  for distinct times  $t_{n+1} > t_n$ . Set  $\lambda := V_{t_n} \cdot \eta(h)$ . Then conditional on  $V_{t_n}$ ,  $V_{t_{n+1}}$  is distributed as  $\exp(-\kappa h)/\eta(h)$  times a non-central chi-squared distribution with  $\nu$  degrees of freedom and non-centrality parameter  $\lambda$ , i.e.

$$\mathbb{P}(V_{t_{n+1}} < x \mid V_{t_n}) = F_{\chi_\nu^2(\lambda)}(x \cdot \eta(h) / \exp(-\kappa h)).$$

Andersen [5] and Haastrecht and Pelsser [22] produce very effective *approximate* sampling methods for the non-central  $\chi_\nu^2(\lambda)$  transition density. Our goal is to compare these methods with the Marsaglia generalized Gaussian (MAGG) method for sampling *exactly* from this transition density outlined in Section 2. Note that for Cox–Ingersoll–Ross sampling using the MAGG, we set  $\lambda = V_{t_n} \cdot \eta(h)$  and compute

$$V_{t_{n+1}} = (Y_1^2 + \cdots + Y_{2N}^2 + Z_1^{2q} + \cdots + Z_p^{2q}) \cdot \exp(-\kappa h) / \eta(h).$$

### 3.2 Asset price process simulation

We follow the lead firstly of Broadie and Kaya [10], and then secondly, of Andersen [5]. By integrating the exact equation for the volatility from  $t_n$  to  $t_{n+1}$  an expression for  $\int \sqrt{V_\tau} dW_\tau^1$  is obtained. This is substituted into the exact equation for  $\ln S_t$ , itself in integral form between  $t_n$  and  $t_{n+1}$ . The result is the exact relation

$$\ln S_{t_{n+1}} = \ln S_{t_n} + \frac{\rho}{\varepsilon}(V_{t_{n+1}} - V_{t_n} - \kappa\theta h) + \left(\frac{\rho\kappa}{\varepsilon} - \frac{1}{2}\right) \int_{t_n}^{t_{n+1}} V_\tau d\tau + \sqrt{1 - \rho^2} \int_{t_n}^{t_{n+1}} \sqrt{V_\tau} dW_\tau^2.$$

Since  $W_t^2$  is independent of the process  $V_t$  we naturally have, in distribution, that

$$\int_{t_n}^{t_{n+1}} \sqrt{V_\tau} dW_\tau^2 \mid \int_{t_n}^{t_{n+1}} V_\tau d\tau = Z \cdot \sqrt{\int_{t_n}^{t_{n+1}} V_\tau d\tau},$$

where  $Z \sim N(0, 1)$ . Now we make the only approximation, suggested by Andersen, to approximate the remaining integral in the expression for  $\ln S_{t_{n+1}}$  by the midpoint rule:

$$\int_{t_n}^{t_{n+1}} V_\tau d\tau \approx \frac{1}{2}h(V_{t_n} + V_{t_{n+1}}).$$

Using these last two replacements, and exponentiating, we arrive at the approximation

$$S_{t_{n+1}} = S_{t_n} \exp(K_0 + K_1 V_{t_n} + K_2 V_{t_{n+1}} + \sqrt{K_3 V_{t_n} + K_4 V_{t_{n+1}}} \cdot Z),$$

where  $K_0 = -h\rho\kappa/\varepsilon$ ,  $K_1 = h(\kappa\rho/\varepsilon - 1/2)/2 - \rho/\varepsilon$ ,  $K_2 = h(\kappa\rho/\varepsilon - 1/2)/2 + \rho/\varepsilon$  and  $K_3 = K_4 = h(1 - \rho^2)/2$ . One caveat remains. The exact process  $S_t$  is a martingale, however using the prescription just given  $\mathbb{E}[S_{t_{n+1}}|S_{t_n}] \neq S_{t_n}$ . However we can correct for this, we quote again from Andersen [5, p. 21].

**Proposition 2** *Let  $K_1, K_2, K_3, K_4$  be defined as above. With  $s := K_2 + \frac{1}{2}K_4$  set*

$$M := \mathbb{E}[\exp(sV_{t_{n+1}})|V_{t_n}] \quad \text{and} \quad K_0^* := -\ln M - (K_1 + \frac{1}{2}K_3)V_{t_n}.$$

*Then if we replace  $K_0$  by  $K_0^*$  in the scheme for  $S_{t_{n+1}}$  above, we have  $\mathbb{E}[S_{t_{n+1}}|S_{t_n}] = S_{t_n}$ .*

Hence the final task is to compute  $M$ . Since we simulate  $V_{t_{n+1}}$  exactly we know  $M = \mathbb{E}[\exp(\hat{s}z)|V_{t_n}]$ , where  $z \sim \chi_\nu^2(\lambda)$ , with  $\nu$  and  $\lambda$  defined for the Heston model, and  $\hat{s} = s \cdot \exp(-\kappa h)/\eta(h)$ . Hence provided  $\hat{s} < \frac{1}{2}$  we have  $M = \exp(\lambda \hat{s}/(1-2\hat{s}))/ (1-2\hat{s})^{\nu/2}$ . Consequently in our simulation for  $S_{t_{n+1}}$ , we take

$$K_0^* = -\frac{\lambda \hat{s}}{1-2\hat{s}} + \frac{1}{2}\nu \ln(1-2\hat{s}) - (K_1 + \frac{1}{2}K_3)V_{t_n}.$$

The requirement  $\hat{s} < \frac{1}{2}$  translates to a mild restriction on the stepsize  $h$ , which in practice is not a problem (see Andersen [5, p. 24]).

	Case I	Case II
$\varepsilon$	1	0.9
$\kappa$	0.5	0.3
$\rho$	-0.9	-0.5
$T$	10	15
$v(0), \theta$	0.04	0.04

**Table 1** Test cases from Andersen. In all cases  $u(0) = 100$ .

### 3.3 Simulation results

We test our Marsaglia generalized Gaussian method (MAGG) directly against Andersen’s QE-M method. The performance of the method of Haastrecht and Pelsser [22] is similar to Andersen’s; the reader interested in the actual comparisons is referred to their paper. We use two of Andersen’s test cases for pricing long-dated European FX call options, maturing at time  $T$  with strike  $K$ ; see Table 1. Let the exact option price at maturity be  $C$ . The error of the approximation is  $E = C - \hat{C}$ , where  $\hat{C}$  is the sample average of the simulated option payout at maturity. In our examples, we use a sample size of  $10^6$ . In Table 2 we show the errors for the two test cases at three different strikes  $K = 100, 140, 60$ ; *without* any postprocessing such as variance reduction. In terms of accuracy the MAGG method competes very favourably with Andersen’s QE-M method, as might be expected. In terms of efficiency, averaging over all stepsizes in case I, the ratio of total CPU times, of the MAGG method over Andersen’s method was 0.4981, making MAGG two times faster. Averaging over all stepsizes in case II, the ratio is 0.8121.

## 4 Concluding remarks

We have seen that the method of sampling from the central  $\chi_\nu^2$  distribution that we have introduced, based on sampling from the generalized Gaussian distribution, is more efficient than the acceptance-rejection method of Ahrens and Dieter. This is true for degrees of freedom  $\nu < 2$  and in particular for values of  $\nu$  given to three or four significant figures. Combining this method for the central  $\chi_\nu^2$  distribution, with a sample from the Poisson distribution with mean  $\lambda/2$ , means that we can efficiently sample from a non-central  $\chi_\nu^2(\lambda)$  distribution, and therefore from the transition probability density of the Cox–Ingersoll–Ross process for the volatility in the Heston model. We demonstrated, that for values of the number of degrees of freedom  $\nu = 4\kappa\theta/\varepsilon^2$  in the Heston model, given to three significant figures, the method we proposed is in fact more efficient than the Andersen method for the same accuracy. An interesting issue that arises here is the structural stability, i.e. sensitivity to the calibrated parameter data, of the Heston model. For example, how sensitive are the computed option prices to having a fourth or fifth significant figure of accuracy? We intend to pursue this question next.

In addition, Glasserman and Kim [20] have recently introduced a novel method for simulating the Heston model. As we can see from our analysis in Section 3.2, to compute the price process at the end-time  $T$ , we in essence need to sample from the distribution for  $\int V_\tau d\tau$  on the interval  $[0, T]$ . The transition density for this integral process over

Stepsize	Case I		Case II	
$h$	Andersen	Marsaglia	Andersen	Marsaglia
Strike 100				
1	0.2211 (0.012)	-0.2374 (0.013)	-0.4833 (0.042)	-0.1404 (0.042)
1/2	0.1164 (0.013)	-0.0707 (0.013)	-0.0400* (0.046)	-0.0264* (0.044)
1/4	0.0143* (0.013)	-0.0440 (0.013)	-0.0231* (0.044)	0.0217* (0.048)
1/8	-0.0277* (0.013)	-0.0050* (0.013)	0.0807* (0.045)	-0.0553* (0.052)
1/16	0.0162* (0.013)	0.0019* (0.013)	-0.0026* (0.042)	0.0521* (0.046)
Strike 140				
1	-0.0883 (0.002)	-0.0283 (0.002)	-0.3082 (0.036)	-0.0926* (0.036)
1/2	-0.0274 (0.003)	-0.0121 (0.002)	0.0515* (0.040)	0.0029* (0.037)
1/4	-0.0013 (0.003)	-0.0048 (0.003)	-0.0016* (0.038)	0.0207* (0.043)
1/8	0.0047 (0.003)	-0.0011 (0.003)	0.0740* (0.039)	-0.0327* (0.047)
1/16	0.0018 (0.003)	0.0015 (0.003)	0.0069* (0.035)	0.0509* (0.040)
Strike 60				
1	0.0317* (0.025)	-0.1234 (0.025)	0.1180 (0.048)	-0.0379* (0.049)
1/2	0.0345* (0.025)	-0.0556* (0.025)	0.1349 (0.052)	-0.0036* (0.050)
1/4	0.0111* (0.025)	-0.0388* (0.025)	-0.0066* (0.050)	0.0290* (0.054)
1/8	0.0407* (0.025)	0.0120* (0.025)	0.0809* (0.052)	-0.0650* (0.058)
1/16	0.0284* (0.025)	0.0003* (0.025)	-0.0170* (0.049)	0.0492* (0.052)

**Table 2** Estimated error using  $10^6$  paths. Sample standard deviations are shown in parenthesis. As Andersen, we star results that are not statistically significant at the level of three sample standard deviations.

the whole interval  $[0, T]$ , given  $V_0$  and  $V_T$ , is well known and given in Pitman and Yor [47]. Its Laplace or Fourier transform has a closed form. Broadie and Kaya [10] use Fourier inversion techniques to sample from this transition density for  $\int V_\tau d\tau$ . Glasserman and Kim instead separate the Laplace transform of this transition density into constituent factors, each of which can be interpreted as the Laplace transforms of probability densities, samples of which can be generated by series of particular gamma random variables. The advantage of this method is that  $\int V_\tau d\tau$  is simulated directly on the interval  $[0, T]$ . However it does require approximation in the form of truncating the series, and the simulation of a sufficiently large number of gamma random variables. It would be of interest to compare the accuracy and efficiency of Glasserman and Kim's approach to ours, especially if we simulate the price process  $S_T$  directly as follows. We decompose  $\int V_\tau d\tau$  on  $[0, T]$  into subintervals  $[t_n, t_{n+1}]$ , use a simple quadrature to approximate  $\int V_\tau d\tau$  on these subintervals much like Andersen, and simulate the transition densities required using the generalized Marsaglia method. We then only exponentiate at the final time  $T$  to generate an approximation for  $S_T$  (since we do not compute the price process at each timestep, this will be more efficient). However, one advantage of the approach we have taken in Section 3.2 for simulating the price process based on the method proposed by Andersen, is that it is more flexible. For example, it allows us to consider pricing path-dependent options. It would also be of interest to compare the accuracy and efficiency of this approach for the pricing of path-dependent options as well.

**Acknowledgements** We would like to thank the referee of a very early version of this manuscript, for encouraging us to perform some numerical comparisons—this lead to Sec-

tion 2.3—and also for pointing out the limit involving the Euler–Mascheroni constant. We also like to thank Gavin Gibson, Mark Owen and Karel in 't Hout for stimulating discussions. A part of this paper was completed while AW was visiting the Fields Institute, Toronto. Their support during the visit is gratefully acknowledged.

## References

1. J.H. AHRENS AND U. DIETER, *Computer generation of Poisson deviates from modified normal distributions*, ACM Transactions on Mathematical Software, 8(2) (1982), pp. 163–179.
2. J.H. AHRENS AND U. DIETER, *Computer methods for sampling from gamma, beta, Poisson, and binomial distributions*, Computing, 12 (1974), pp. 223–246.
3. A. ALFONSI, *On the discretization schemes for the CIR (and Bessel squared) processes*, Monte Carlo Methods and Applications, 11(4) (2005), pp. 355–384.
4. A. ALFONSI, *A second-order discretization scheme for the CIR process: application to the Heston model*, Preprint, 2007.
5. L. ANDERSEN, *Simple and efficient simulation of the Heston stochastic volatility model*, Journal of Computational Finance, 11(3) (2008), pp. 1–42.
6. L.B.G. ANDERSEN AND V.V. PITERBARG, *Moment explosions in stochastic volatility models*, Finance and Stochastics, 11(1) (2007), pp. 29–50.
7. A. BERKAOUI, M. BOSSY AND A. DIOP, *Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence*, ESAIM Probability and Statistics, 12 (2008), pp. 1–11.
8. P. BILLINGSLEY, *Probability and measure*, Wiley series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., 1995.
9. M. BOSSY AND A. DIOP, *An efficient discretization scheme for one dimensional SDEs with a diffusion coefficient function of the form  $|x|^\alpha$ ,  $\alpha \in [1/2, 1)$* , INRIA Rapport de recherche n° 5396, 2007.
10. M. BROADIE AND Ö. KAYA, *Exact simulation of stochastic volatility and other affine jump diffusion processes*, Operations Research, 54(2) (2006), pp. 217–231.
11. P. CARR AND D.B. MADAN, *Option valuation using the fast Fourier transform*, Journal of Computational Finance, 2(4) (1999), pp. 61–73.
12. G.D. CHALMERS AND D.J. HIGHAM, *First and second moment reversion for a discretized square root process with jumps*, Preprint December 17, 2008.
13. J.C. COX, J.E. INGERSOLL AND S.A. ROSS, *A theory of the term structure of interest rates*, Econometrica, 53(2) (1985), pp. 385–407.
14. G. DEELSTRA AND F. DELBAEN, *Convergence of discretized stochastic (interest rate) processes with stochastic drift term*, Appl. Stochastic Models Data Anal., 14 (1998), pp. 77–84.
15. S. DYRTING, *Evaluating the non-central chi-square distribution for the Cox–Ingersoll–Ross process*, Computational economics, 24 (2004), pp. 35–50.
16. F. FANG AND K. OOSTERLEE, *Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions*, MPRA Paper No. 9248, posted 20. June 2008, online at [texttthttp://mpa.ub.uni-muenchen.de/9248](http://mpa.ub.uni-muenchen.de/9248)
17. F. FANG AND K. OOSTERLEE, *A novel pricing method for European options based on Fourier-cosine series expansions*, SIAM J. Sci. Comput., 30(2) (2008), pp. 826–848.
18. W. FELLER, *Two singular diffusion problems*, Annals of Mathematics, 54 (1951), pp. 173–182.
19. P. GLASSERMAN, *Monte Carlo methods in financial engineering*, Applications of Mathematics, Stochastic Modelling and Applied Probability, 53, Springer, 2003.
20. P. GLASSERMAN AND K. KIM, *Gamma expansion of the Heston stochastic volatility model*, Finance Stoch., DOI 10.1007/s00780-009-0115-y.
21. A.K. GUPTA AND D. SONG,  *$L_p$ -norm spherical distribution*, Journal of Statistical Planning and Inference, 60 (1997), pp. 241–260.
22. A. VAN HAASTRECHT AND A. PELSSER, *Efficient, almost exact simulation of the Heston stochastic volatility model*, Preprint November 17th, 2008.
23. S.L. HESTON, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Review of Financial Studies, 6(2) (1993), pp. 327–343.
24. D. HOBSON, *Stochastic volatility models, correlation, and the  $q$ -optimal measure*, Mathematical Finance, 14 (2004), pp. 537–556.

25. D.J. HIGHAM AND X. MAO, *Convergence of Monte-Carlo simulations involving the mean-reverting square root process*, Journal of Computational Finance, 8 (2005), pp. 35–61.
26. D.J. HIGHAM, X. MAO AND A.M. STUART, *Strong Convergence of Numerical Methods for Nonlinear Stochastic Differential Equations*, SIAM J. Num. Anal, 40 (2002), pp. 1041–1063.
27. K. IN 'T HOUT, J. BIERKENS, A.P.C. VAN DER PLOEG AND J. IN 'T PANHUIS, *A semi closed-form analytic pricing formula for call options in a hybrid Heston–Hull–White model*, Proceedings of the Study Group Mathematics with Industry 2007, University of Utrecht, 2007.
28. A. JENTZEN, P.E. KLOEDEN AND A. NEUENKIRCH, *Pathwise approximation of stochastic differential equations on domains: Higher order convergence rates without global Lipschitz coefficients*, Numerische Mathematik, 112(1) (2009), pp 41–64.
29. N.L. JOHNSON, *On an extension of the connection between Poisson and  $\chi^2$  distributions*, Biometrika, 46 (3/4) (1959), pp. 352–363.
30. C. KAHL, M. GÜNTHER AND T. ROSSBERG, *Structure preserving stochastic integration schemes in interest rate derivative modeling*, Appl. Numer. Math., 58(3) (2008), pp. 284–295.
31. C. KAHL AND P. JÄCKEL, *Fast strong approximation Monte Carlo schemes for stochastic volatility models*, Quant. Finance, 6(6) (2006), pp. 513–536.
32. C. KAHL AND P. JÄCKEL, *Not-so-complex logarithms in the Heston Model*, Wilmott, September 2005, pp. 94–103.
33. C. KAHL AND H. SCHURZ, *Balanced Milstein methods for ordinary SDEs*, Monte Carlo Methods Appl., 12(2) (2006), pp. 143–170.
34. I. KARATZAS AND S.E. SHREVE, *Brownian motion and stochastic calculus*, Graduate texts in mathematics, 2nd Edition, Springer, 1988.
35. S. KARLIN AND H.M. TAYLOR, *A second course in stochastic processes*, Academic Press, 1981.
36. D.E. KNUTH, *The art of computer programming, vol 2: Seminumerical algorithms*, Addison-Wesley, Reading, Mass., Third edition, 1998.
37. A. LIPTON AND A. SEPP, *Stochastic volatility models and Kelvin waves*, J. Phys. A: Math. Theor., 41 (2008), pp. 1–23.
38. G. LORD, S. J. A. MALHAM AND A. WIESE, *Efficient strong integrators for linear stochastic systems*, SIAM J. Numer. Anal., 46(6) (2008), pp. 2892–2919.
39. R. LORD, R. KOEKKOEK AND D. VAN DIJK, *A comparison of biased simulation schemes for stochastic volatility models*, Tinbergen Institute Discussion Paper TI2006-046/4, 2006.
40. S.J.A. MALHAM AND A. WIESE, *Stochastic Lie group integrators*, SIAM J. Sci. Comput., 30(2) (2008), pp. 597–617.
41. S.J.A. MALHAM AND A. WIESE, *Stochastic expansions and Hopf algebras*, Proc. R. Soc. A, 465 (2009), pp. 3729–3749.
42. G. MARSAGLIA, *Improving the polar method for generating a pair of random variables*, Boeing Sci. Res. Lab., D1-82-0203, 1962.
43. C.D. MEYER, *Matrix analysis and applied linear algebra*, SIAM, 2000.
44. E. MORO AND H. SCHURZ, *Boundary preserving semianalytical numerical algorithms for stochastic differential equations*, SIAM J. Sci. Comput., 29(4) (2007), pp. 1525–1549.
45. S. NINOMIYA AND N. VICTOIR, *Weak approximation of stochastic differential equations and application to derivative pricing*, arXiv:math/0605361v3, 1 Oct 2006.
46. P.B. PATNAIK, *The non-central  $\chi^2$ - and F-distributions and their applications*, Biometrika, 36 (1949), pp. 202–232.
47. J. PITMAN, AND M. YOR, *A decomposition of Bessel bridges*, Z. Wahrscheinlichkeitstheor. Verw. Geb., 59 (1982), pp. 425–457.
48. D. REVUZ AND M. YOR, *Continuous Martingales and Brownian motion*, Springer-Verlag, 1991.
49. H. SCHURZ, *A brief introduction to numerical analysis of (ordinary) stochastic differential equations without tears*, in Handbook of Stochastic Analysis and Applications, V. Lakshmikantham and D. Kannan, eds., Marcel Dekker, 2002, pp. 237–359.
50. A. SEPP, *Pricing options on realized variance in the Heston model with jumps in returns and volatility*, The Journal of Computational Finance, 11(4) (2008), pp. 33–70.
51. F. SINZ AND M. BETHGE, *How much can orientation selectivity and contrast gain control reduce redundancies in natural images*, Max-Planck-Institute für Biologische Kybernetik, Technical Report, March 2008.



- 
52. F. SINZ, S. GERWINN AND M. BETHGE *Characterization of the  $p$ -generalized normal distribution*, Journal of Multivariate Analysis, 100 (2009), pp. 817–820.
  53. R.D. SMITH, *An almost exact simulation method for the Heston model*, Journal of Computational Finance, 11(1) (2007), pp. 115–125.
  54. D. SONG AND A.K. GUPTA,  *$L_p$ -norm uniform distribution*, Proceedings of the AMS, 125(2) (1997), pp. 595–601.
  55. J. WILKIE AND Y.M. WONG, *Positivity preserving chemical Langevin equations*, Chemical Physics, 353 (2008), pp. 132–138.

